PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) **International Patent Classification:** **G06F 7/58, H04L 9/20**	**A1**	(11) **International Publication Number:**	**WO 00/25203**
		(43) **International Publication Date:**	04 May 2000 (04.05.2000)

(21) **International Application Number:** PCT/SG99/00105

(22) **International Filing Date:** 26 October 1999 (26.10.1999)

(30) **Priority Data:**
9803458-0 28 October 1998 (28.10.1998) SG

(60) **Parent Application or Grant**
DATAMARK TECHNOLOGIES PTE LTD. [/]; (). HO, Anthony, Tung, Shuen [/]; (). TAM, Siu, Chung [/]; (). TAN, Siong, Chai [/]; (). YAP, Lian, Teck [/]; (). HO, Anthony, Tung, Shuen [/]; (). TAM, Siu, Chung [/]; (). TAN, Siong, Chai [/]; (). YAP, Lian, Teck [/]; (). NAMAZIE, Farah ; ().

Published

(54) **Title: METHODS OF DIGITAL STEGANOGRAPHY FOR MULTIMEDIA DATA**

(54) **Titre: PROCEDES DE STEGANOGRAPHIE NUMERIQUE DESTINES A DES DONNEES MULTIMEDIA**

(57) Abstract

A lossless steganographic encoding method for secure transmission or storage of multimedia data. Primary data, such as text, image, video, audio or other digital data, is utilised in a steganographic process to encode secondary data, such as text, image, video, audio or other digital data. The primary data includes a plurality of first data elements and the secondary data includes a plurality of second data elements. For each second data element an operation is performed with a first data element so as to generate a key element as a result of the operation. The key elements may then be securely transmitted and/or stored. In preferred embodiments of the method, the primary data may be rearranged according to a predefined or random manner, or it may be resized so as to match the size of the secondary data. A complementary decoding method is disclosed, and a method of generating a pseudo-random number sequence, which may be used in the steganographic and decoding methods, is also disclosed.

(57) Abrégé

La présente invention concerne un procédé de codage stéganographique sans perte qui permet de protéger la transmission ou le stockage de données multimédia. Des données primaires telles que du texte, des images, des données vidéo, des données audio ou d'autres données numériques sont utilisées dans un procédé stéganographique pour coder des données secondaires telles que du texte, des images, des données vidéo, des données audio ou d'autres données numériques. Les données primaires comprennent une pluralité de premiers éléments de données et les données secondaires comprennent une pluralité de deuxièmes éléments de données. Pour chaque deuxième élément de données une opération est effectuée avec un premier élément de données de manière à générer un élément clé représentant le résultat de l'opération. Les éléments clés peuvent ensuite être transmis et/ou stockés de manière protégée. Dans des formes de réalisation préférées de l'invention, les données primaires peuvent être réarrangées suivant une manière prédéfinie ou aléatoire, ou bien être redimensionnées afin de correspondre à la taille des données secondaires. On décrit un procédé de décodage complémentaire ainsi qu'un procédé permettant de générer une séquence de nombres pseudo-aléatoires qui peut être utilisé dans des procédés de codage et de décodage stéganographiques.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G06F 7/58, H04L 9/20 — A1

(11) International Publication Number: WO 00/25203

(43) International Publication Date: 4 May 2000 (04.05.00)

(21) International Application Number: PCT/SG99/00105

(22) International Filing Date: 26 October 1999 (26.10.99)

(30) Priority Data:
9803458-0 28 October 1998 (28.10.98) SG

(71) Applicant *(for all designated States except US)*: DATAMARK TECHNOLOGIES PTE LTD. [SG/SG]; Suite 106, Innovation Centre, Block 1, 16 Nanyang Drive, Singapore 637722 (SG).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: HO, Anthony, Tung, Shuen [CA/SG]; 54H Nanyang View #09-16, Singapore 639669 (SG). TAM, Siu, Chung [SG/SG]; 78B Eng Kong Place, Singapore 599154 (SG). TAN, Siong, Chai [SG/SG]; Block 426, Fajar Road #01-545, Singapore 670426 (SG). YAP, Lian, Teck [SG/SG]; Block 312, 32 Bukit Batok Street #11-79, Singapore 650312 (SG).

(74) Agents: NAMAZIE, Farah et al.; Haq & Namazie Partnership, Robinson Road, P.O. Box 765, Singapore 901515 (SG).

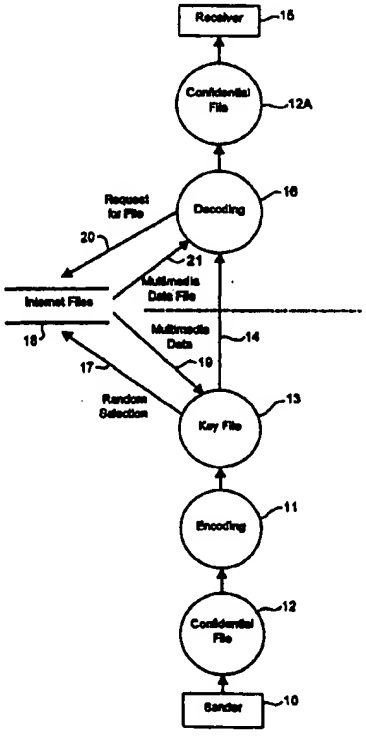
(81) Designated States: AU, CA, CN, ID, JP, KR, SG, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
With amended claims.

(54) Title: METHODS OF DIGITAL STEGANOGRAPHY FOR MULTIMEDIA DATA

(57) Abstract

A lossless steganographic encoding method for secure transmission or storage of multimedia data. Primary data, such as text, image, video, audio or other digital data, is utilised in a steganographic process to encode secondary data, such as text, image, video, audio or other digital data. The primary data includes a plurality of first data elements and the secondary data includes a plurality of second data elements. For each second data element an operation is performed with a first data element so as to generate a key element as a result of the operation. The key elements may then be securely transmitted and/or stored. In preferred embodiments of the method, the primary data may be rearranged according to a predefined or random manner, or it may be resized so as to match the size of the secondary data. A complementary decoding method is disclosed, and a method of generating a pseudo-random number sequence, which may be used in the steganographic and decoding methods, is also disclosed.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Description

METHODS OF DIGITAL STEGANOGRAPHY FOR MULTIMEDIA DATA

Field of the Invention

The present invention relates generally to steganographic methods of encoding digital data for secure transmission or storage of information. The invention also relates to complementary decoding methods and to a method of generating a pseudo-random number sequence using any digital file. The pseudo-random number sequence may be used in the steganographic encoding or decoding methods.

The encoding method is especially suited to digital camouflaging or steganography for confidential information such as text, audio, still image or video data, and it will be convenient to describe the method in relation to that example application. It should be appreciated, however, that the encoding method is intended for broader application and use. Similarly, the method of generating a pseudo-random number sequence may be used in applications other than steganography applications.

Background of the Invention

The tremendous growth in multimedia products and services provided via the Internet and digital data storage media (DSM) has led to the need for copyright authentication and for protecting data integrity. In the past few years, a number of digital watermarking techniques have been developed for the purpose of resolving legal use issues associated with copyright information on the Internet and DSM.

A number of digital watermarking techniques have recently been patented. Examples of these include US Patent 5,636,292 to Rhoads (1997) and US Patent 5,659,726 to Sandford and Handel (1997). Rhoads discloses methods to impress an identification code on a carrier, such as an electronic data signal or a physical medium, in a manner that permits the identification code to be later discerned and the carrier thereby identified. Sandford and Handel disclose a method of embedding auxiliary information into host data, such as a photograph,

television signal, facsimile transmission, or identification card. The method operates by manipulating a noise component of the host data in accordance with the auxiliary information.

Many prior art digital watermarking techniques, including the techniques disclosed in the above US patents, are only able to conceal limited information, such as a few logical bits (ie. "1" and "0") or a few characters (eg. "A12"), in the host data. However, to record detailed ownership information for a host work in which copyright subsists, such as a satellite image of Singapore, an entire message or sentence may need to be concealed in, or associated with, the host data. For example, the sentence "Digital Image of Singapore is the property of Mr John Tan, dated 16 December 1997" may provide more conclusive proof as to true ownership of the host work than having to rely on just a simple code to assess copyright infringement.

There therefore remains a need for a steganographic encoding method which may allow a relatively long string of secondary data (such as text, image, audio or video data) to be encoded using primary data (such as text, image, audio or video data) without degradation of the primary data.

Besides the above mentioned application on the Internet, many potential consumer, commercial and service applications may benefit from the use of digital steganography technology, including for copyright protection and signature authentication purposes and for secure transmission of information. These applications include steganographic encoding of secured text, image, audio or video data containing ownership identification or attribute information associated with digital still or video cameras, copyright protection and royalty tracking of sound recordings in the music industry. Commercial and service sectors may also benefit from secure transmission and reception of confidential information and digital signature associated with sensitive documents and electronic transactions that could be encoded in normal data streams transmitted through an open channel.

Pseudo-random number generators are algorithms or devices that give a fixed sequence of random numbers when the seed is the same. This seed may be a number, a bit-stream, a digital file or any other form of data.

Typical random number generators use hashing functions for example, SHA (secure hash algorithm), as in US Patent Number 5,787,179 awarded to Microsoft Corporation (1998), and US Patent Number 5,732,138 awarded to Silicon Graphics Inc. (1998).

Summary of the Invention

In one aspect, the present invention provides a method of generating a pseudo-random number sequence including the steps of:

providing source data including an ordered plurality of data elements, the content of each data element being represented by a group of digits;

reading the groups of digits into an array such that each position in the array contains one of said digits;

selecting a starting position within the array of digits; and

regrouping said digits to form new groups of digits with reference to the starting position, such that each new group represents a pseudo-random number and successive new groups represent said pseudo-random number sequence.

In one embodiment the data elements of the source data are represented in binary notation and the content of each data element is preferably represented by a byte (ie. 8 bits). In this embodiment, each bit of each 8-bit byte constitutes a digit which may be read into a bit array such that each position in the array contains one bit.

The starting position may be selected randomly, pseudo-randomly or in a pre-defined manner. Based on that starting position the bits are regrouped into new groups of preferably eight bits, each new group constituting a new byte of information. In this way, each new byte represents a pseudo-random number which bears no numerical relationship to numerical values of the data elements of the source data.

The term "pre-defined" as used throughout this specification refers to that which is defined or can be defined by a user or by the program.

The source data may be obtained from a digital file available in the public domain, a private database, or any digital storage medium (DSM). The file may

represent a text sequence, an image, an audio sequence, a video sequence, a graphics representation, a computer program, or any accessible digital data.

Unlike the abovementioned prior art random number generators which use a hashing function, the present invention uses the whole or part of a digital file. The contents of digital files can be considered as random depending on the location selected for the starting position and how the bits are grouped. As a result, the same digital file with different starting positions and grouping methods will generate completely different pseudo-random number sequences. Different digital files with the same starting position and the same grouping method will also generate completely different pseudo-random number sequences. This has the distinct advantage that it is able to regenerate the same sequence of pseudo-random numbers as long as the same digital file, the same starting position, and the same grouping method are used. Since this method is not based on any mathematical formula, there is no way of obtaining the same sequence of random numbers without knowing the source file, the starting position, and the grouping method.

Advantageously, the pseudo-random number sequence is stored for use in a steganographic data encoding or decoding method, a cryptographic encoding or decoding method, or for any other purpose requiring a sequence of random numbers.

In another aspect, the present invention provides an encoding method including the steps of:

providing primary data including an ordered plurality of first data elements;

providing secondary data including a plurality of second data elements; and

for each second data element

(i) performing an operation with a first data element, and

(ii) generating a key element as a result of said operation.

In one embodiment the encoding method includes, prior to performing said operations, a step of rearranging the first data elements of the primary data. A plurality of techniques for rearranging the first data elements may be available and a selection may be made from the plurality of techniques. The selection may

be made randomly or pseudo-randomly, or by a user. The first data elements may be rearranged in a predefined manner or in a random or pseudo-random manner. Alternatively, or additionally, similar rearranging steps may be performed on the second data elements of the secondary data.

In one embodiment the primary data is in the form of a primary data array containing the first data elements and the secondary data is in the form of a secondary data array containing the second data elements. The encoding method may include a step of resizing the primary data array to match the size of the secondary data array. If the secondary data array is smaller than the primary data array, the primary data array may be truncated to match the size of the secondary data array. If the secondary data array is larger than the primary data array, first data elements of the primary data array may be repeated so as to increase the size of the first data array to match that of the secondary data array. In an embodiment including a rearranging step as well as a resizing step, the repeated first data elements may be rearranged according to techniques other than the technique selected for rearranging the first group of first data elements. In other words, although the first data elements of the primary data may be multiplied, each group of multiplied first data elements need not necessarily be rearranged according to the same technique as the first group of first data elements. Moreover, each repeated group may be arranged according to a different technique.

The operation to be performed between the first and second data elements may include a mathematical operation, a logical operation, a mapping function, or any other operation which serves to generate key elements as a result of the operation. Preferably, a plurality of operations is available and a selection is made from the plurality of operations. The selection may be made randomly or pseudo-randomly, or by a user.

The encoding method may generate a string of key elements which is associated with a corresponding string of second data elements. Unique key data, which is generated for given primary and secondary data, may be stored for use in a complementary decoding method, as described below.

Preferably the key elements are stored in a key file, which may then be

transmitted or archived for future use. Advantageously, information about the encoding process, such as the operation performed, the rearranging technique, etc., is also stored in the key file. This information may be stored within a header or attribute section of the key file. An attribute section may be positioned anywhere in the key file, not necessarily at the beginning.

The source, primary, secondary and key data mentioned above may be represented in digital binary form. However, any form of data representation or notation, using any convenient set of symbols, may be used, eg. alphanumeric characters, integer numbers, etc. The primary data may represent or be derived from a still image, motion video, audio, text or other type of information. Likewise, the secondary data may represent a still image, motion video, audio, text or other information.

In a preferred form of the invention, the secondary data includes a text message and each second data element includes an alphanumeric character. However, each secondary data element may include a character from another character set. The alphanumeric characters may be used to compose the text message. In a typical application of the invention the text message may include confidential information relating to an image, a video or an audio sequence contained in the primary data. In one embodiment, the text message may include one or more of the following: a title, an artist, a copyright holder, a body to which royalties should be paid, and general terms of publisher distribution.

In other embodiments, the text message may be a confidential message, a representation of an image, a representation of an audio sequence, or a combination of the above.

The primary data may represent a text message, a still image, an audio sequence, a motion video segment, general multimedia data, a graphics file, a complete program, or any other accessible digital data that can be retrieved from the public domain, such as an Internet website, a private database, the random access memory or buffer of a computer, or any digital storage medium. The first data elements of the primary data may be arranged in an array.

Each first data element may define a characteristic associated with a still image element. The first data elements may be obtained from a stream of data

representing a digitised still image. The image may be obtained from an Internet web site, a digital camera, a computer game, computer software or other source. It may be a greyscale or color image (wherein each first data element defines a grey level or colour component, for example) and may be stored in any known format, eg. BMP, GIF, TIFF, or JPEG.

Alternatively, or additionally, each first data element may define a characteristic associated with a motion video element. The first data elements may be obtained from a stream of data representing digitised motion video. The digitised video may be obtained from an Internet web site, a Video Compact Disc (VCD) player, a Laser Disc (LD) player, a computer game, computer software, a Digital Versatile Disc (DVD) player or other source, and may be stored in any known format, eg. MPEG or AVI.

Alternatively, or additionally, each first data element may define a characteristic associated with a digital audio sample. The digital audio samples may be obtained from a stream of data representing digitised sound or music. The digitised sound may be obtained from an Internet web site, a Compact Disc (CD) player, Digital Audio Tape (DAT) player, Laser Disc player, Video Compact Disc (VCD) player or other source, and may be stored in any known format eg. WAV, AIFF, MIDI, etc. In one embodiment, the digital audio samples are obtained from two streams of data representing two channels of digitised sound for stereo reproduction.

In the preferred embodiment of the encoding method, the primary data includes a random or pseudo-random number sequence. The still image, motion video or audio data mentioned in the preceding three paragraphs may be used as source data for generating a pseudo-random number sequence according to the method described above. That number sequence, based on the original image, video or audio data, may then be used as primary data in the encoding method of the invention.

In an alternative embodiment, the primary data may be obtained from a conventional random-number generator or other suitable source.

In another aspect, the present invention provides a method of decoding secondary data including a plurality of second data elements, said secondary

data being encoded in a plurality of key elements such that each key element is generated by an operation performed with a respective first data element of primary data, said method including the steps of:

providing said primary data including an ordered plurality of said first data elements;

providing said plurality of key elements;

for each key element, generating a corresponding said second data element by performing an inverse of said operation.

Compared with existing steganographic or digital watermarking techniques the present invention has the distinct advantage that long sentences of text, large amounts of data of any form, e.g. images, audio, video, or any binary files, may be encoded and subsequently decoded in confidence. With any form of data, e.g. images, audio, video, binary files, digital bit patterns, the integrity of the primary data is never affected or compromised in any way. As such, the primary data may be transmitted by any means e.g. by mail, e-mail, telephone, fax, ftp, http, dial-up networking, local area network, wide area network, Internet, Intranet, Extranet, or by any other electronic means. The data can also be retrieved from any storage medium, such as hard disk, floppy disk, zip disk, CD ROM, DAT, VCD, DVD. In a preferred way, since the primary data is never modified, there is no need to re-send the primary data for every message. Only the key data has to be sent. Therefore, this method results in lower bandwidth usage and faster transmission via a communication channel when compared to any existing steganographic or watermarking technique.

In an alternative embodiment, when access to open or stored data, eg. Internet, CD ROM, VCD or DVD, etc., is restricted or limited at the receiving end of the transmission channel, the primary or source data (in whole or in part) may also be sent as part of the key file. This embodiment of the invention offers a lower level of security but may be preferred by some users for its convenience. To improve security in this embodiment, a password or other protection may be implemented in conjunction with the invention. This embodiment of the invention can then form part of a larger system for transmitting confidential information.

In a modified version of the latter embodiment, the primary or source data

(in whole or in part) may be sent as a separate file with proper identification.

Brief Description of the Drawings

The accompanying drawings, which are incorporated into and constitute part of the description of the invention, illustrate embodiments of the invention and serve to explain the principles thereof. It is to be understood, however, that the drawings and following detailed description are given for the purposes of illustration only and are not intended as a definition of the limits of the invention.

In the drawings:

Figure 1 shows a context diagram showing an example application of the invention for confidential data transmission;

Figure 2 shows a flow-chart of a preferred embodiment of the invention incorporating a two-part steganographic encoding method;

Figure 3 shows an example of rearranging a primary data file for use in the steganographic encoding method;

Figure 4 shows an example of a mathematical operation;

Figure 5 shows an example of a logical XOR operation between primary and secondary data;

Figure 6 shows an example of a 1:1 mapping operation; and

Figure 7 shows an example of the steganographic encoding method performed on a password.

Description of Preferred Embodiments

A preferred embodiment of the invention uses source or primary data, such as a still image, motion video, audio, text or other data, to steganographically encode secondary data, such as a data file containing confidential information. The confidential information may likewise include a still image, motion video, audio, text or any other type of data. The encoding process generates unique key data representing the secondary data in an encoded form.

One embodiment of the invention, to be described in detail below, includes two main processes. The first main process uses source data, such as a still

image, motion video, audio, text or other data, to generate an array containing a pseudo-random number sequence. That array of pseudo-random numbers is then used as primary data in a second main process to steganographically encode the secondary data.

The source data may be provided as a file containing the image, video, audio, text or other data. For ease of description, this file will be referred to as the Container File. Similarly, the secondary data may be provided as a file which, for ease of description, will be referred to as the Confidential File. The key data may also be stored to a file, which will be referred to as the Key File.

Referring now to Figure 1, there is shown a preferred embodiment of the invention used for secure transmission of confidential data over an open communication channel. The sender 10 performs a steganographic encoding process 11 on a Confidential File 12 so as to generate a unique Key File 13 which may be securely transmitted over the open communication channel 14. The receiver 15 of the Key File 13 performs a complementary decoding process 16 on that file to retrieve the Confidential File 12A.

To steganographically encode the Confidential File 12, either the sender 10 or the encoding process 11 selects 17 from the Internet 18 a data file to be downloaded 19 for use as the Container File in the encoding process 11. After performing the encoding process 11 and generating the Key File 13, the sender 10 can transmit the Key File 13 to the receiver 15 over the open channel 14. The receiver 15 can then send a request 20 to the Internet 18 to download 21 the same Container File at his/her end and perform the decoding process 16 on the Key File 13.

The sender 10 and receiver 15 may have agreed on a particular Internet file to use as the Container File in the encoding and decoding processes. Alternatively, the Key File 13 may carry information on where to find the Internet file used by the sender.

As mentioned above, the Container File and Confidential File may contain any types of data. Accordingly, one can choose to encode a video file using an audio file, an image file using a text file, or any other combination. The invention does not constrain the user to a particular combination.

Referring now to Figure 2, there is shown a flowchart illustrating in more detail the two-part staganographic encoding process of the preferred embodiment of the invention. Steps 30-32 relate to the first main process for generating an array of pseudo-random numbers based on source data (Container File) and steps 33-37 relate to the second main process of steganographically encoding secondary data (Confidential File) using the array of pseudo-random numbers as primary data to generate key data (Key File).

Main Process 1

This process generates an array of pseudo-random numbers based on a source file containing digital data.

In step 30, a digital source file (Container File) containing a plurality of bytes of data is read into an array of bits. The source file may be any type of file containing any type of information, eg. audio, video, image, text, etc.

In step 31, one of the elements of the bit array is selected as a starting position. This selection may be made in a random or pseudo-random manner or in a predefined manner.

In step 32, the elements of the bit array are regrouped into new groups of bytes (8 bits) beginning from the starting position. In this manner, the resulting new groups represent pseudo-random numbers in a sequence which may be stored as an array.

It should be appreciated that this process is applicable to number systems other than one based on two (ie. binary). That is, the digital information carried in the source data need not necessarily be converted into bits. If the information is converted into a decimal system, or a number system with a base of 16, etc., the same principle may be applied to create new random numbers.

The regrouping step performed in step 32 need not always regroup the bits into new groups of eight. Supposing the binary system is used, and the array of bits is regrouped into bytes, the range of the generated random numbers will be from 0 to 255. If instead the bits are regrouped into nibbles (4 bits), the range will be narrower (0-15). For a larger range, the groups can be made

larger. For other number base systems, the size of the groups chosen may similarly be varied.

Because this process is not confined to any particular medium, the user has a very large number of files to choose from and use as the Container File. Even when the same file is used, the possibilities for selecting a starting position are numerous. The flexibility of the process allows the user to generate many possible random number arrays. It can therefore serve as a good tool for formatting the source data file prior to steganographically encoding a secondary data file. In other words, the process described above is a preferred preliminary process to apply before applying Main Process 2, described below.

Main Process 2

This process steganographically encodes secondary data (Confidential File) using primary data (eg. the array of pseudo-random numbers obtained from Step 32 in Main Process 1) to generate key data (Key File). Alternatively, the primary data may be obtained from a conventional random number generator or from an image, video, audio, text, or other digital data file.

In step 33 of Figure 2, the primary data array of random numbers is rearranged so as to increase the difficulty of breaking the code. The user may be provided with a wide choice of techniques for rearranging the array of random numbers so as to further increase the difficulty of hacking. The selection of the rearranging technique may be determined randomly. For example, a password may be used as a seed to generate a pseudo-random number (for example by the use of the RAND() function in the C programming language) to select a rearrangement technique. Alternatively, the user may be allowed to define or select the rearrangement technique to apply.

The technique of rearranging may be in a predefined or pseudo-random manner. Examples include: arranging in the reverse order, scanning row-by-row, column-by-column, in a zig-zag manner, or in a spiral manner, etc. Figure 3 shows an example of rearranging a typical data stream from a Container File 38 in the reverse order 39. As a further example, the spiral method involves first

taking the element at the X position, then the element at the (X+1) position, then the element at the (X-1) position, then the (X+2) position, then the (X-2) position, and so on.

The rearranging step is optional and may be omitted if it is felt that the degree of randomness introduced by applying a random number generator to the source data file is sufficient. In the preferred embodiment, the random number array is rearranged to introduce a higher degree of randomness.

In Step 34 the primary data array of random numbers may be resized to match the size of the secondary data array of second data elements contained in the Confidential File. The array of random numbers may be larger or smaller than the array of secondary data. The array of random numbers is therefore either truncated or repeated so as to match the size of the array of secondary data array. Therefore, whether this step is necessary depends on the relative sizes of the arrays and on the types of operations performed or to be performed in subsequent steps of the process.

In the event that the secondary data array is larger than the array of random numbers, all or part of the array of random numbers is repeated and the repeated random numbers may be rearranged (in Step 33) according to a different technique. In this manner, more random numbers may be provided for the subsequent operation in Step 35, described below.

In Step 35, at least one operation is performed between elements of the array of random numbers and elements of the secondary data array contained in the Confidential File. This results in a key array which contains the results of the operations.

Because each operation is between at least one random number and at least one element of the secondary data, the result obtained is different even for similar elements of the secondary data. For example, given an array of random numbers [3, 5, 2,....] and an array of second data elements [1, 3, 1,....], and supposing the operation chosen is to subtract the values of the second data elements from the random numbers, the key array obtained will be [2, 2, 1,....]. The first and third elements of the secondary data array are identical but produce different key elements because of the way in which the random numbers are

utilised in the encoding process. This is an important advantage of the invention because it makes cracking of the code more difficult.

Furthermore, the invention does not limit the user to the selection of the operation(s) to perform, thus making hacking even more difficult.

Various types of operations may be performed, including the following:

(i) A mathematical operation such as subtraction. An example of such an operation is shown in Figure 4 wherein second data elements 40 of the Confidential File are subtracted from first data elements 41 of the random number array to generate key elements 42. Other mathematical operations may include addition, multiplication, etc.

(ii) A logical operation, such as the XOR operation. Such an operation is shown in Figure 5 wherein each bit of each second data element 50 is XORed with a corresponding bit of each first data element 51 to generate a resultant bit of each key element 52.

(iii) A 1:1 mapping function. An example of such a function is illustrated in Figure 6 wherein mapping is based on the index positions as specified by the second data elements. For example, if the content of a second data element 60 has a value of "2", then "2" is taken as an index pointing to the random number 61 at position 2. The random number 61 at position 2 has a value of "98" and this is taken to be the value to be stored in the corresponding key element 62 of the key array.

The selection of operation(s) to be performed may be determined randomly. For example, a password may be used as a seed to generate a pseudo-random number (for example by the use of the RAND() function in C) to choose an operation to be performed. Alternatively, the user may be allowed to define or select the operation(s) to perform.

Referring again to Step 35 of Figure 2, the results of the operation are stored in a key array. In Step 36, information about the encoding process is stored in a header or attribute file, which is then combined in Step 37 with the key array to form a Key File. The Information Header or Attribute Section of the Key File contains all necessary information to perform the complementary decoding process. Such information may include the physical location of the Container

File, the starting position for the pseudo-random number generation process, the techniques and means of rearranging the array of random numbers, the operation performed, etc.

The encoding process may optionally include a password feature to increase security. The sender may provide a password which is also put through the encoding process. At the other end, the receiver may be prompted to enter a password and decoding is performed on the encoded password provided by the sender. Only if the decoded password matches that provided by the receiver will the decoding process proceed to reproduce the Confidential File. This process is illustrated in Figure 7 wherein a Password Array 70 containing the password "HelloWorld" is represented by the ASCII code 72, 101, 108, etc. These ASCII codes are then subtracted from the random numbers 71 to create key elements 72. These key elements are then stored in the attribute section of the Key File.

It should be understood that the data transmission application shown in Figure 1 may or may not incorporate the two-part encoding process shown in Figure 2. For example, the first main process for generating the pseudo-random number array on the Container File may be omitted. In that event, the Container File may be used as primary data in the encoding process instead of the random number array.

Further, it should be understood that the rearranging and resizing steps within the encoding process, Main Process 2, are optional and may be omitted.

It is considered that the complementary decoding process would be self evident to those skilled in the art from the information presented herein. The decoding process need not therefore be described in detail. Clearly, a key part of the decoding process is to perform an inverse operation of that performed in the encoding process. If rearranging and resizing of the primary data (ie. the random number array) has been performed in the encoding process, details must be stored in the attribute section of the Key File, or elsewhere, so that a reverse operation may be performed during the decoding process. Similarly, if a random number array has been generated from a source data file using Main Process 1, that same random number array must again be reproduced from the source data file for use in decoding of the Key File.

Advantages of the invention

A) Unrestricted secondary data size

Compared with existing steganographic or watermarking techniques the present invention has the distinct advantage that long sentences of text, large amount of data of any form, e.g. images, audio, video, binary files, may be encoded (camouflaged) and subsequently decoded in confidence.

B) No distortion in primary data or secondary data

With any form of data, e.g. images, audio, video, binary files, digital bits patterns etc., the integrity of the primary data or secondary data is never affected or compromised in any way. In other words, the decoding technique is lossless. The primary data may be optionally transmitted in any form e.g. by mail, telephone, e-mail, fax, ftp, http, dial-up networking, local area network, wide area network, Internet, intranet, or by any other electronic means. The data can also be retrieved from any storage medium, such as hard disk, floppy disk, zip disk, DAT, CD, VCD, LD, DVD. This invention has a significant advantage over the conventional methods, such as least significant bit (LSB) coding, which impose distortion to the data, thus the whole Container File must be sent. Apart from that, LSB coding allows only high bit-depth Container Files to be used, thus it is not applicable to most multimedia data..

C) Lower bandwidth usage and faster transmission

In a preferred way, since the primary data is never modified, there is no need to send or re-send the primary data for every message. Only the Key File needs to be sent. This results in reduced storage space used compared with conventional methods which require the whole Container File to be sent. Therefore, this method results in a lower bandwidth usage and faster transmission down a communication channel compared to any existing steganographic or watermarking technique.

D) Unrestricted primary data type and secondary data type

Existing steganographic and watermarking techniques usually have problems with low bit-depth bitmaps (e.g. black & white images), low bit-depth

audio and video files. This is usually due to the problem that altering the least significant bit of low bit-depth files would change the original information too much. This restricts existing steganographic or watermarking techniques to be applicable only to large bit-depth files, such as a 24-bit bitmaps, etc. However, since the present invention maintains the integrity of both the primary data and secondary data, it does not suffer from this problem and thus is able to be used for any primary data type or secondary data type.

E) Unique generated key data

The invention disclosed above has another distinct advantage in that even with the same primary data and secondary data, the generated key data is always different and unique. This makes it almost impossible for any hacker to crack the code by analysing the generated key data.

F) Different rearranging techniques

Many rearranging techniques may be implemented in this invention. This means that hackers must attempt all the rearranging techniques in order to break the code. Given that hacking a single technique is already an extremely difficult task, breaking the code becomes virtually impossible.

G) Unlimited primary data available

With the tremendous growth in Internet communication, the number of primary data files available on the Internet is practically infinite. Thus, intended users can select an image, audio, video or any digital binary file on the Internet to be used as the primary data. Thus, without the knowledge of the primary data, hackers have to try an infinite number of images, audio and video files before they can proceed with the hacking mission.

H) Password protection and a garbage-in-garbage-out system

This invention includes a garbage-in-garbage-out password protection system. The password may be used to generate the random rearranging method and/or the starting location of the primary data and or secondary data to start. Since this is designed as a garbage-in-garbage-out system, it does not give any clue as to whether the password is invalid or the primary data is invalid. Therefore, even if hackers manage to get information on the primary data file, which is already very difficult, constantly hacking the key file with various

passwords without any success may finally lead the hackers to think that the primary data file is not the right one.

l) Generation of new Container Files

Unique primary data files known only to the intended users can be easily generated. Examples of these could be a digital image of the intended users, an audio speech of the intended users, and a video clip of the intended users.

Typical Applications of the Invention

In one embodiment, the invention may be used for confidential data communication. In a preferred way, the primary data may be predetermined and the generated key file may then be transmitted to the intended users e.g. by mail, telephone, video conferencing, e-mail, fax, ftp, http, dial-up networking, Internet, Intranet, or by any other electronic means. It is found that the size of the Key File that needs to be sent is almost of equal size to the actual message, with an overhead usually of fewer than 10 bytes.

In another embodiment, the invention may be implemented as a plug-in for an Internet web browser, e-mail program, graphics program, document program or any other computer program so that confidential data can be hidden and sent only to intended users.

In yet another embodiment, software developers who want to protect their data can also apply the invention disclosed above. For example, in Microsoft® Word, the program can use the password and the document itself to hide the original data. Only the user who is able to enter the correct password would be able to view the document. Therefore, even if other programs are able to open Microsoft® Word documents, the opened document will still be presented as unintelligent data. In the same manner, this embodiment may be extended to other programs for example, an e-mail program such as Exchange™, or a graphic software such as AutoCAD®.

In a further embodiment, the invention may be used as a data verifier for the detection of modification of a sent message. The sent message in this case may be considered as the primary data while a digital signature of the sender

may be considered as the secondary data or vice versa. Upon receiving the message, the receiver can decode it to detect if the actual sender has sent it and to check if that message has been modified.

In another embodiment, confidential information or authentication codes may be stored in credit cards, passports, identity cards, cash cards, or any devices in which both primary data and secondary data exist. For example, in the case of credit cards, the biometrics (eg. photographs, fingerprints, voice, etc.) of the credit card owner may be used as the primary data while the information about the owner or his/her account or the authentication codes may be considered as the secondary data or vice versa. In such a case, if an attempt were made to change the biometrics of the credit card owner, the decoded confidential information or authentication codes would not tally.

In another embodiment, the technique may be used to generate a digital watermark in any digital image, text, audio, video or any other digital data. The image, text, audio, video or digital data may be considered as the primary data (Container File) while the digital watermark may be considered as the secondary data (Confidential File). In the encoding method, a Key File will be generated according to the invention disclosed. The rightful owner will hold the unique Key File and he can use it to decode the digital watermark from the primary data, thence proving the originality of the primary data.

In another embodiment, part of the current invention (Main Process 1 described above) may be used in the field of cryptography. In cryptography, no container file is used as in the case of steganography. Instead, a hashing function is used to decode an encrypted message. This hashing function may be a password string or a very large prime number known only to the sender and the receiver. Therefore, the pseudo-random number sequence generated using Main Process 1 can be used in place of any hashing function. Again, in view of the many possible digital files available in both the public and private domains, and the ease of making new digital files, hacking the pseudo-random number sequence will be extremely difficult if not impossible.

In yet another embodiment, the current invention may also be applied complementarily to the field of cryptography. Using the current invention, either

the hashing function or the encrypted message may be encoded and subsequently decoded for added security. Alternatively, the Key File generated using the current invention may be encrypted before transmission to the sender for subsequent decryption before being decoded steganographically.

It is anticipated that the invention will be modelled and implemented in software on general-purpose computer platforms. Alternatively, the invention may be implemented using hardwired circuitry, CPU, DSP and incorporated in one or more application specific ICs. Further, it is anticipated that the invention may be embedded into facsimile machines, telephones, digital cameras, walkie-talkies or other electronic messaging devices to enable the encoding and decoding of confidential information.

Finally, those skilled in the art will appreciate that various adaptations and modifications of the just described preferred embodiments may be configured without departing from the scope and the spirit of the invention. Therefore, it is to be understood that within the scope of the appended claims, the invention may be practiced other than as specifically described herein.

Claims

CLAIMS

1. A method of generating a pseudo-random number sequence including the steps of:

providing source data including an ordered plurality of data elements, the content of each data element being represented by a group of digits;

reading the groups of digits into an array such that each position in the array contains one of said digits;

selecting a starting position within the array of digits; and

regrouping said digits to form new groups of digits with reference to the starting position such that each new group represents a pseudo-random number and successive new groups represent said pseudo-random number sequence.

2. A method according to claim 1, further including the step of storing said pseudo-random number sequence.

3. A method according to claim 1 wherein the data elements are represented in binary notation.

4. A method according to claim 3 wherein each new group of digits includes eight binary digits.

5. A method according to claim 1 wherein the starting position is selected randomly or pseudo-randomly.

6. A method according to claim 1 wherein the starting position is selected in a pre-defined manner.

7. An encoding method utilising the pseudo-random number sequence generated by a method according to claim 1.

8. A decoding method utilising the pseudo-random number sequence

generated by a method according to claim 1.

9. An encoding method including the steps of:

providing primary data including an ordered plurality of first data elements;

providing secondary data including a plurality of second data elements; and

for each second data element

(i) performing an operation with a first data element, and

(ii) generating a key element as a result of said operation.

10. An encoding method according to claim 9 including, prior to performing said operations, the step of:

rearranging the first data elements of the primary data.

11. An encoding method according to claim 10 wherein a plurality of techniques for rearranging the first data elements is available and at least one selection is made from the plurality of techniques.

12. An encoding method according to claim 11 wherein the or each selection is made randomly or pseudo-randomly.

13. An encoding method according to claim 11 wherein the or each selection is made in a pre-defined manner.

14. An encoding method according to claim 11 including the steps of:

storing the key elements in a key file; and

storing information about the or each selected rearranging technique in an attribute section of the key file.

15. An encoding method according to claim 10 wherein the first data elements are rearranged in a predefined manner.

16. An encoding method according to claim 10 wherein the first data elements are rearranged in a random or pseudo-random manner.

17. An encoding method according to claim 9 including, prior to performing said operations, the step of:

rearranging the second data elements of the secondary data.

18. An encoding method according to claim 9 wherein the primary data is in the form of a primary data array containing the first data elements and the secondary data is in the form of a secondary data array containing the second data elements, further including the step of:

resizing the primary data array to match the size of the secondary data array.

19. An encoding method according to claim 18 wherein resizing includes the step of:

if the secondary data array is smaller than the primary data array, truncating the primary data array, and

if the secondary data array is larger than the primary data array, repeating first data elements of the primary data array.

20. An encoding method according to claim 19 including, prior to performing said operations, the step of rearranging the first data elements of the primary data array according to a first technique, and rearranging the repeated first data elements according to said first technique or further techniques other than said first technique.

21. An encoding method according to claim 9 wherein the first and second data elements are represented by numbers and wherein each operation includes a mathematical operation between the first and second data elements.

22. An encoding method according to claim 9 wherein the first and second

data elements are represented in binary notation and each operation includes a logical operation between the first and second data elements.

23. An encoding method according to claim 9 wherein the first and second data elements are represented by numbers and each operation is a mapping function.

24. An encoding method according to claim 9 wherein the first and second data elements are represented by numbers and each operation is a 1:1 mapping function wherein the content of each second data element is used as an index for selecting a first data element and the content of each selected first data element is assigned to the associated key element.

25. An encoding method according to claim 9 wherein a plurality of operations is available and a selection is made from the plurality of operations.

26. An encoding method according to claim 25 wherein the selection is made randomly or pseudo-randomly.

27. An encoding method according to claim 25 wherein the selection is made in a pre-defined manner.

28. An encoding method according to claim 9 including the step of storing the key elements in a key file.

29. An encoding method according to claim 28 including the step of storing information about the encoding process within an attribute section of the key file.

30. An encoding method according to claim 29 wherein the information stored in the attribute section includes the operation or operations performed.

31. An encoding method according to claim 28 including the step of storing

the primary data in the key file.

32. An encoding method according to claim 9 wherein the primary data includes the pseudo-random number sequence generated by a method according to claim 1.

33. An encoding method according to claim 9 wherein the primary data includes a random number sequence generated by a random number generator.

34. An encoding method according to claim 9 wherein the primary data is provided from a file obtained from the Internet.

35. An encoding method according to claim 34 including the steps of:

storing the key elements in a key file; and

storing information about the Internet file in an attribute section of the key file.

36. An encoding method according to claim 9 wherein the secondary data includes a text message and each second data element includes a character from a character set.

37. An encoding method according to claim 9 wherein the first data elements are arranged in an array and each first data element represents a characteristic associated with a digital audio sample.

38. An encoding method according to claim 9 wherein the first data elements are arranged in an array and each first data element represents a characteristic associated with a still image element.

39. An encoding method according to claim 9 wherein the first data elements are arranged in an array and each first data element represents a characteristic associated with a motion video element.

40. A method of decoding secondary data including a plurality of second data elements, said secondary data being encoded in a plurality of key elements such that each key element is generated by an operation performed with a respective first data element of primary data, said method including the steps of:

providing said primary data including an ordered plurality of said first data elements;

providing said plurality of key elements;

for each key element, generating a corresponding said second data element by performing an inverse of said operation.

41. A method according to claim 40 wherein during encoding of the secondary data, the first data elements are rearranged according to a defined technique prior to performing the operations, said method including, prior to generating said second data elements, the step of:

rearranging the first data elements of the primary data according to said defined technique.

42. A method according to claim 41 wherein the key elements are provided in a key file having an attribute section and the attribute section contains information about said defined technique for rearranging the first data elements during the encoding of the secondary data, said method including the step of reading said information from the attribute section for determining said defined technique.

43. A method according to claim 40 wherein during encoding of the secondary data, the primary data is resized to match the size of the secondary data, said method including, prior to generating said second data elements, the step of resizing the primary data according to the resizing performed during the encoding of the secondary data.

44. A method according to claim 43 wherein during encoding of the secondary

data, the primary data is resized by truncating the primary data if the secondary data is smaller than the primary data or by repeating the primary data if the secondary data is larger than the primary data, said method including, prior to generating the second data elements, the step of:

if the primary data was truncated during encoding, truncating the primary data according to the truncating performed during the encoding of the secondary data; and

if the primary data was repeated during encoding, repeating the primary data according to the repeating performed during the encoding of the secondary data.

45. A method according to claim 44 wherein during encoding of the secondary data, the first data elements of the primary data are rearranged according to a first technique and repeated first data elements are rearranged according to said first technique or further techniques other than said first technique, said method including, prior to generating said second data elements, the step of rearranging the first data elements of the primary data array according to said first technique, and rearranging the repeated first data elements according to said first technique or said further techniques.

46. A method according to claim 40 wherein the key elements are provided in a key file having an attribute section and the attribute section contains information about the operations performed during the encoding of the secondary data, said method including the step of reading said information from the attribute section for determining for each key element said inverse of said operation.

47. A method according to claim 40 wherein during encoding of the secondary data, the primary data is provided from a file obtained from the Internet, and the key elements are provided in a key file having an attribute section which contains information about the Internet file, said method including the step of reading said information from the attribute section for retrieving said Internet file.

48. A method according to claim 40 wherein the primary data includes a pseudo-random number sequence generated by a method according to claim 1.

AMENDED CLAIMS

[received by the International Burcau on 08 April 2000 (08.04.00) ;
original claims 1-48 replaced by new claims 1-45 (8 pages)]

CLAIMS

1. An encoding method including steps of:
providing primary data including an ordered plurality of first data elements;
providing secondary data including a plurality of second data elements; and
for each second data element
(i) performing an operation with a first data element, and
(ii) generating a key element as a result of said operation;
wherein each operation is performed and each key element is generated without degrading said primary data.

2. An encoding method according to claim 1 including, prior to performing said operations, a step of:
rearranging the first data elements of the primary data.

3. An encoding method according to claim 2 wherein a plurality of techniques for rearranging the first data elements is available and at least one selection is made from the plurality of techniques.

4. An encoding method according to claim 3 wherein the or each selection is made randomly or pseudo-randomly.

5. An encoding method according to claim 3 wherein the or each selection is made by a user.

6. An encoding method according to claim 3 including steps of:
storing the key elements in a key file; and
storing information about the or each selected rearranging technique in an attribute section of the key file.

7. An encoding method according to claim 2 wherein the first data elements are rearranged in a predefined manner.

8. An encoding method according to claim 2 wherein the first data elements are rearranged in a random or pseudo-random manner.

9. An encoding method according to claim 1 including, prior to performing said operations, a step of:

rearranging the second data elements of the secondary data.

10. An encoding method according to claim 1 wherein the primary data is in the form of a primary data array containing the first data elements and the secondary data is in the form of a secondary data array containing the second data elements, further including a step of:

resizing the primary data array to match the size of the secondary data array.

11. An encoding method according to claim 10 wherein resizing includes a step of:

if the secondary data array is smaller than the primary data array, truncating the primary data array, and

if the secondary data array is larger than the primary data array, repeating first data elements of the primary data array.

12. An encoding method according to claim 11 including, prior to performing said operations, a step of rearranging the first data elements of the primary data array according to a first technique, and rearranging the repeated first data elements according to said first technique or further techniques other than said first technique.

13. An encoding method according to claim 1 wherein the first and second data elements are represented by numbers and wherein each operation includes a mathematical operation between the first and second data elements.

14. An encoding method according to claim 1 wherein the first and second data elements are represented in binary notation and each operation includes a logical operation between the first and second data elements.

15. An encoding method according to claim 1 wherein the first and second data elements are represented by numbers and each operation is a mapping function.

16. An encoding method according to claim 1 wherein the first and second data elements are represented by numbers and each operation is a 1:1 mapping function wherein the content of each second data element is used as an index for selecting a first data element and the content of each selected first data element is assigned to the associated key element.

17. An encoding method according to claim 1 wherein a plurality of operations is available and a selection is made from the plurality of operations.

18. An encoding method according to claim 17 wherein the selection is made randomly or pseudo-randomly.

19. An encoding method according to claim 17 wherein the selection is made by a user.

20. An encoding method according to claim 1 including a step of storing the key elements in a key file.

21. An encoding method according to claim 20 including a step of storing information about the encoding process within an attribute section of the key file.

22. An encoding method according to claim 21 wherein the information stored in the attribute section includes the operation or operations performed.

23. An encoding method according to claim 20 including a step of storing the primary data in the key file.

24. An encoding method according to claim 1 wherein the primary data includes a pseudo-random number sequence generated by a method including steps of:

providing said ordered plurality of first data elements, the content of each data element being represented by a group of digits;

reading the groups of digits into an array such that each position in the array contains one of said digits;

selecting a starting position within the array of digits; and

regrouping said digits to form new groups of digits with reference to the starting position such that each new group represents a pseudo-random number and successive new groups represent said pseudo-random number sequence.

25. A method according to claim 24, wherein said method for generating said pseudo-random number sequence includes a step of storing said pseudo-random number sequence.

26. A method according to claim 24 wherein the data elements are represented in binary notation.

27. A method according to claim 26 wherein each new group of digits includes eight binary digits.

28. A method according to claim 24 wherein the starting position is selected randomly or pseudo-randomly.

29. A method according to claim 24 wherein the starting position is selected in a pre-defined manner.

30. An encoding method according to claim 1 wherein the primary data includes a random number sequence generated by a random number generator.

31. An encoding method according to claim 1 wherein the primary data is provided from a file obtained from the Internet.

32. An encoding method according to claim 31 including steps of:
storing the key elements in a key file; and
storing information about the Internet file in an attribute section of the key file.

33. An encoding method according to claim 1 wherein the secondary data includes a text message and each second data element includes a character from a character set.

34. An encoding method according to claim 1 wherein the first data elements are arranged in an array and each first data element represents a characteristic associated with a digital audio sample.

35. An encoding method according to claim 1 wherein the first data elements are arranged in an array and each first data element represents a characteristic associated with a still image element.

36. An encoding method according to claim 1 wherein the first data elements are arranged in an array and each first data element represents a characteristic associated with a motion video element.

37. A method of decoding secondary data including a plurality of second data elements, said secondary data being encoded in a plurality of key elements generated by an operation performed with a respective first data element of primary data, wherein each operation is formed and each key element is generated without degrading said primary data, said method including steps of:

providing said primary data including an ordered plurality of said first data elements;

providing said plurality of key elements; and

for each key element, generating a corresponding said second data element by performing an inverse of said operation.

38. A method according to claim 37 wherein during encoding of the secondary data, the first data elements are rearranged according to a defined technique prior to performing the operations, said method including, prior to generating said second data elements, a step of:

rearranging the first data elements of the primary data according to said defined technique.

39. A method according to claim 38 wherein the key elements are provided in a key file having an attribute section and the attribute section contains information about said defined technique for rearranging the first data elements during the encoding of the secondary data, said method including a step of

reading said information from the attribute section for determining said defined technique.

40. A method according to claim 37 wherein during encoding of the secondary data, the primary data is resized to match the size of the secondary data, said method including, prior to generating said second data elements, a step of resizing the primary data according to the resizing performed during the encoding of the secondary data.

41. A method according to claim 40 wherein during encoding of the secondary data, the primary data is resized by truncating the primary data if the secondary data is smaller than the primary data or by repeating the primary data if the secondary data is larger than the primary data, said method including, prior to generating the second data elements, a step of:

if the primary data was truncated during encoding, truncating the primary data according to the truncating performed during the encoding of the secondary data; and

if the primary data was repeated during encoding, repeating the primary data according to the repeating performed during the encoding of the secondary data.

42. A method according to claim 41 wherein during encoding of the secondary data, the first data elements of the primary data are rearranged according to a first technique and repeated first data elements are rearranged according to said first technique or further techniques other than said first technique, said method including, prior to generating said second data elements, a step of rearranging the first data elements of the primary data array according to said first technique, and rearranging the repeated first data elements according to said first technique or said further techniques.

43. A method according to claim 37 wherein the key elements are provided in a key file having an attribute section and the attribute section contains information about the operations performed during the encoding of the secondary data, said method including a step of reading said information from the attribute section for determining for each key element said inverse of said operation.

44. A method according to claim 37 wherein during encoding of the secondary data, the primary data is provided from a file obtained from the Internet, and the key elements are provided in a key file having an attribute section which contains information about the Internet file, said method including a step of reading said information from the attribute section for retrieving said Internet file.

45. A method according to claim 37 wherein the primary data includes a pseudo-random number sequence generated by a method including steps of:

providing said ordered plurality of first data elements, the content of each data element being represented by a group of digits;

reading the groups of digits into an array such that each position in the array contains one of said digits;

selecting a starting position within the array of digits; and

regrouping said digits to form new groups of digits with reference to the starting position such that each new group represents a pseudo-random number and successive new groups represent said pseudo-random number sequence.

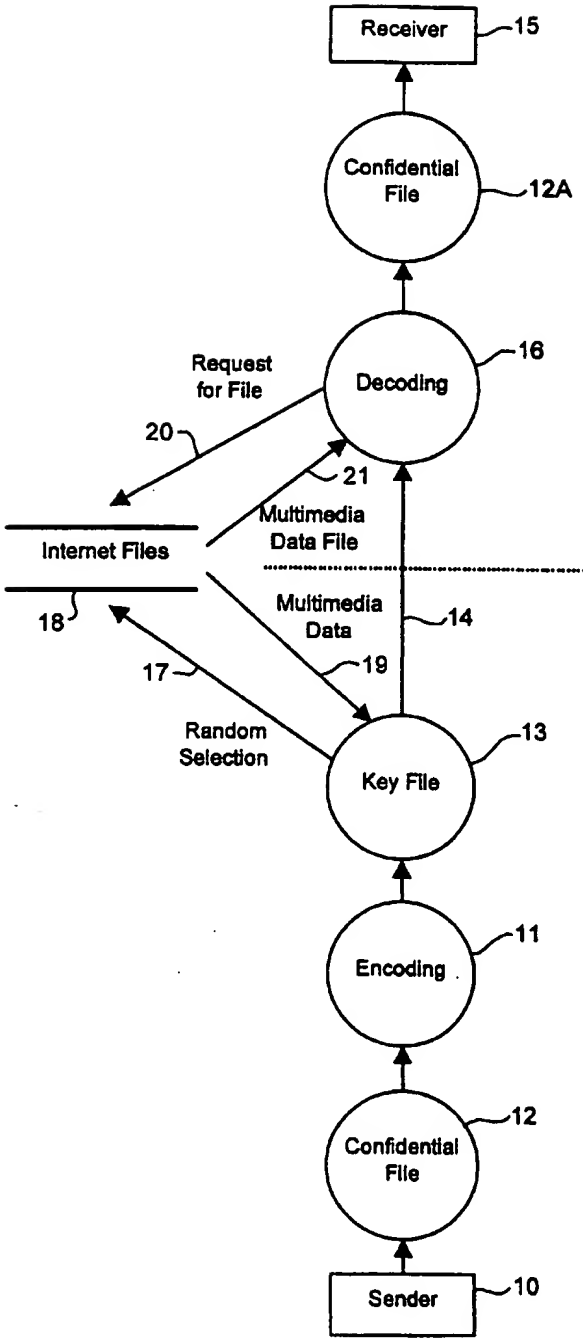

Receiver
15
Confidential File
12A
Decoding
16
Request for File
20
21
Multimedia Data File
Internet Files
18
Multimedia Data
14
19
17
Random Selection
Key File
13
Encoding
11
Confidential File
12
Sender
10

FIG 1

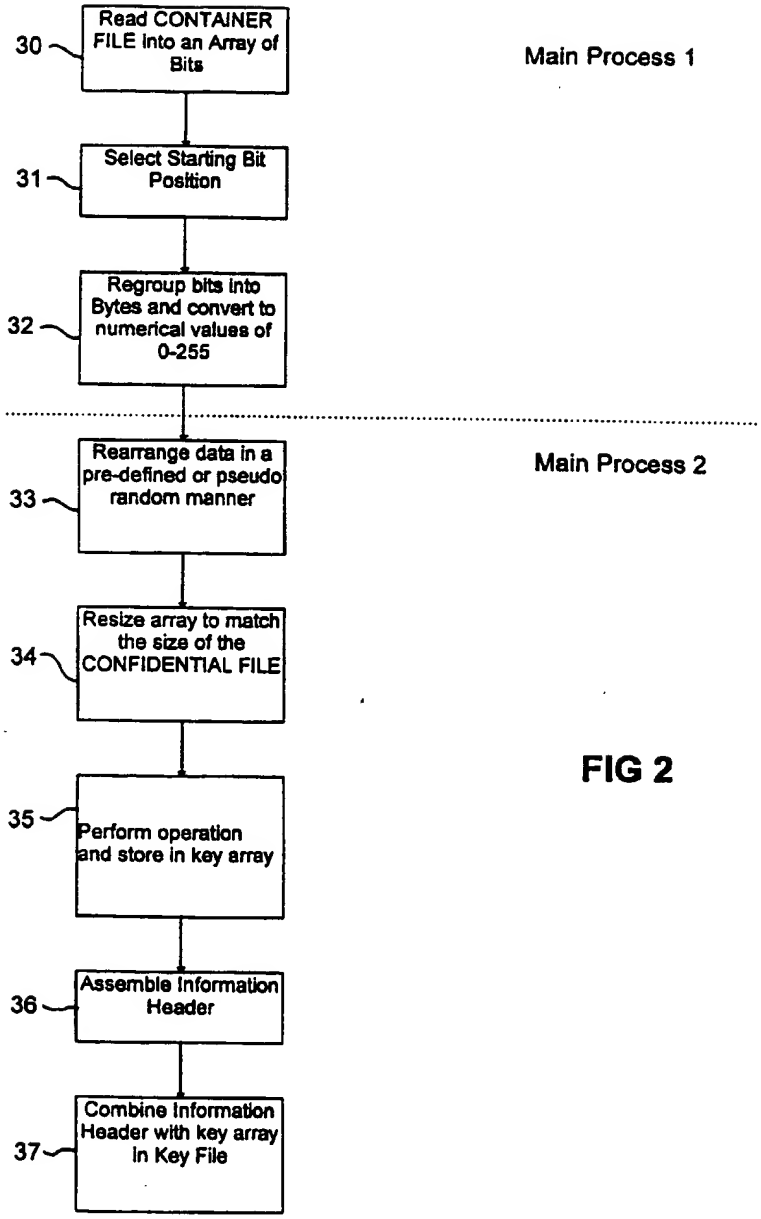

30
Read CONTAINER FILE into an Array of Bits
31
Select Starting Bit Position
32
Regroup bits into Bytes and convert to numerical values of 0-255
Main Process 1
33
Rearrange data in a pre-defined or pseudo random manner
34
Resize array to match the size of the CONFIDENTIAL FILE
35
Perform operation and store in key array
36
Assemble Information Header
37
Combine Information Header with key array in Key File
Main Process 2

FIG 2

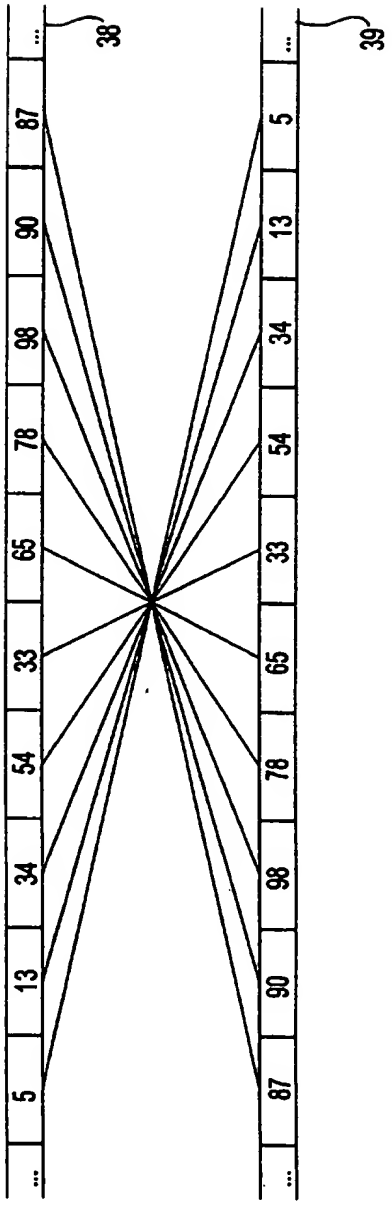


FIG 3

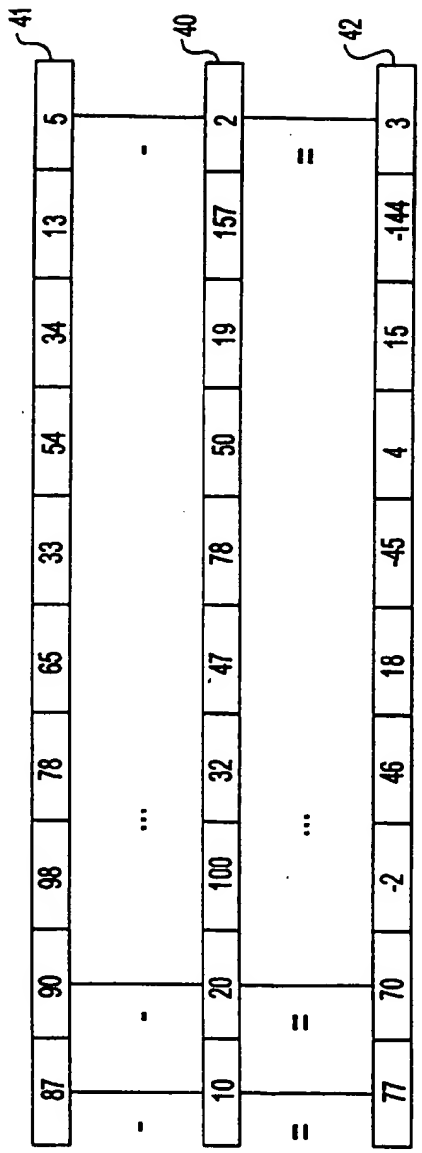

41
40
42
5
13
34
54
33
65
78
98
90
87
2
157
19
50
78
47
32
100
20
10
3
-144
15
4
-45
18
46
-2
70
77
-
=

FIG 4

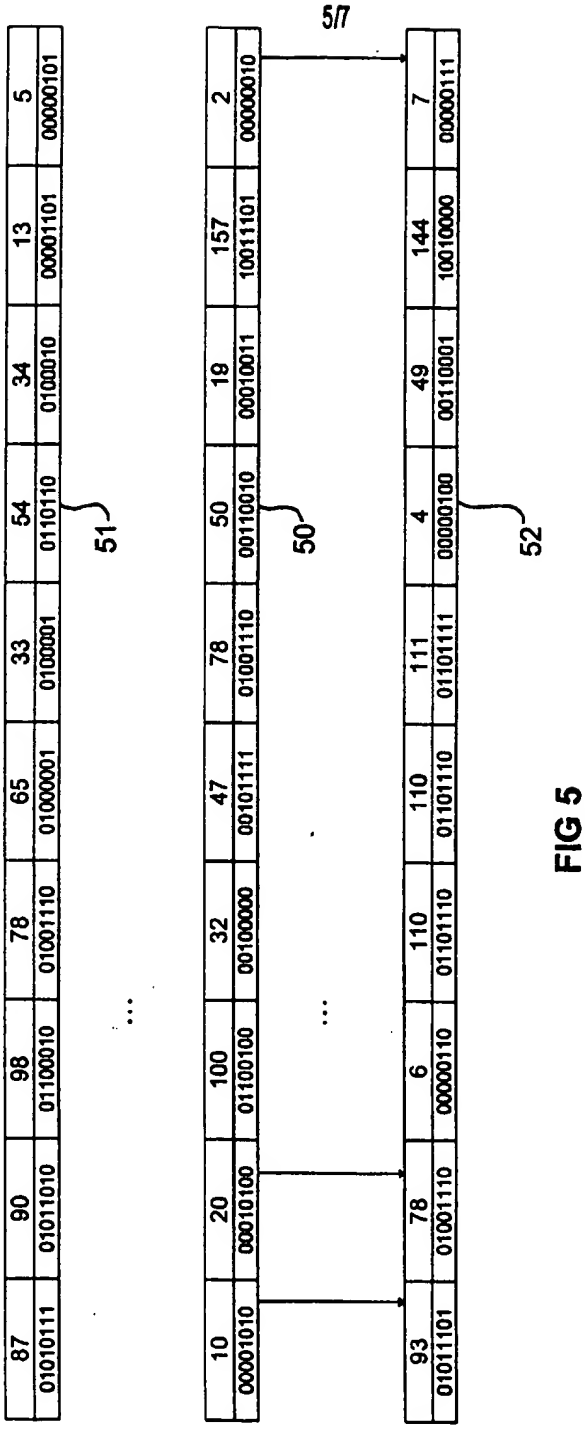

87 90 98 78 65 33 54 34 13 5
01010111 01011010 01100010 01001110 01000001 01000001 01101110 01000010 00001101 00000101
51
10 20 100 32 47 78 50 19 157 2
00001010 00010100 01100100 00100000 00101111 01001110 00110010 00010011 10011101 00000010
50
93 78 6 110 110 111 4 49 144 7
01011101 01001110 00000110 01101110 01101110 01101111 00000100 00110001 10010000 00000111
52

FIG 5

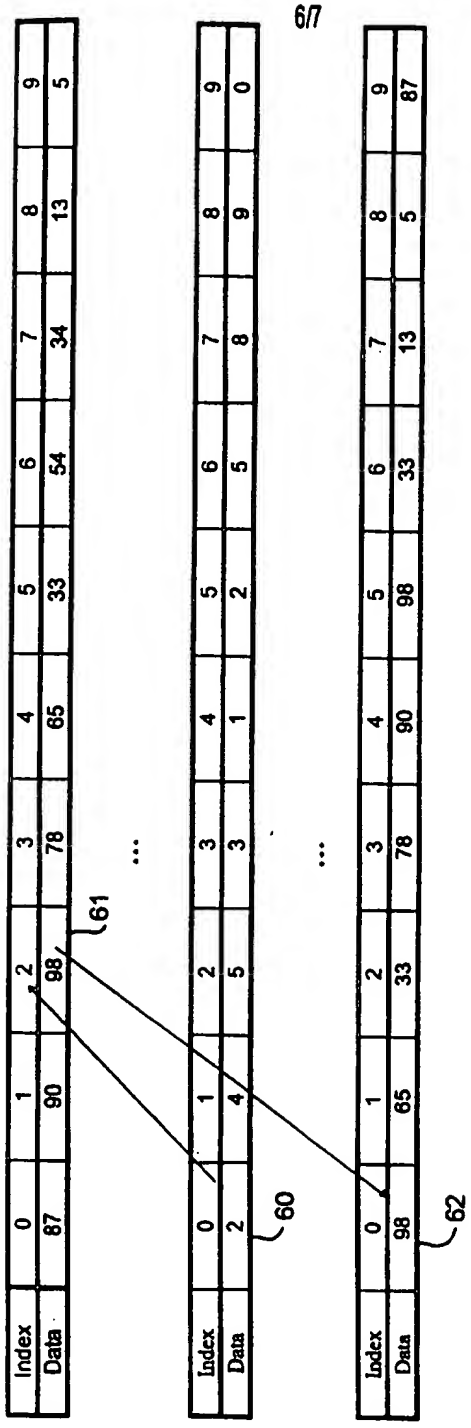


FIG 6

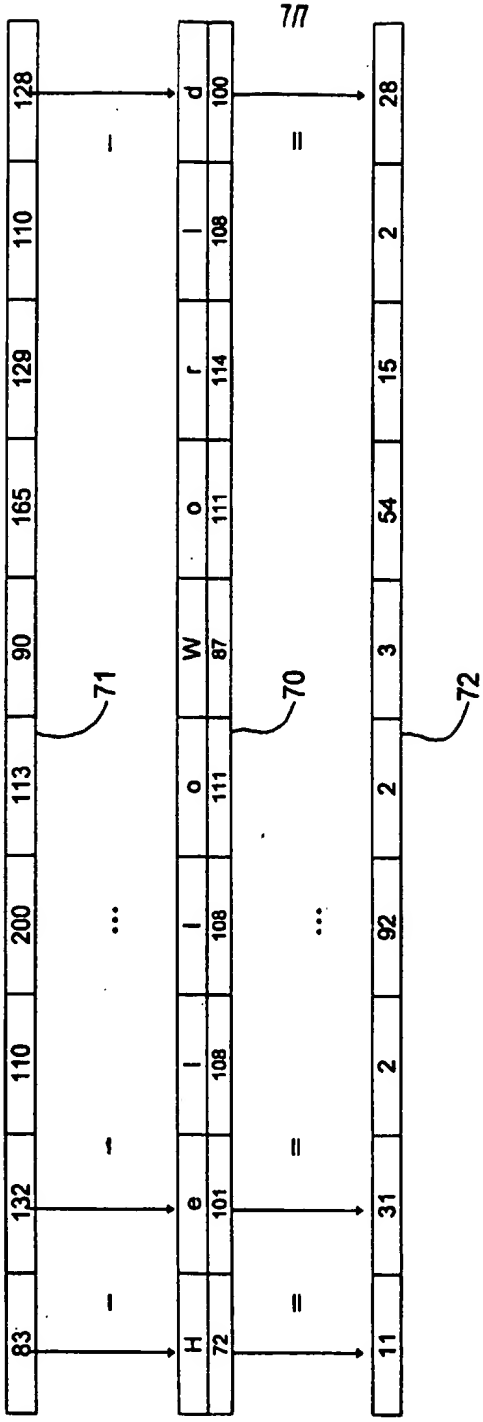

77
128
110
129
165
90
113
200
110
132
83
d
l
r
o
W
o
l
l
e
H
100
108
114
111
87
111
108
108
101
72
28
2
15
54
3
2
92
2
31
11
71
70
72
–
=

FIG 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SG 99/00105

A. CLASSIFICATION OF SUBJECT MATTER

Int Cl[6]: G06F 7/58, H04L 9/20

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC G06F 7/-, H04L 9/-

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
WPAT

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96/42151 A (THE DICE COMPANY) 27 December 1996 pages 14-19	9-11, 17, 21-23, 25, 37-41, 43
A	US 5276738 A (HIRSCH) 4 June 1994 Whole document	9-48
A	EP 301383 A (ADVANTEST CORPORATION) 19 July 1988 Whole document	1-8

☐ Further documents are listed in the continuation of Box C

☒ See patent family annex

* Special categories of cited documents:

"A" Document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 January 2000	11 FEB 2000

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE
PO BOX 200
WODEN ACT 2606 AUSTRALIA
E-mail address: pct@ipaustralia.gov.au
Facsimile No.: (02) 6285 3929

Authorized officer

J. LAW
Telephone No.: (02) 6283 2179

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/SG 99/00105

Box I **Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box II **Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1. Claims 1-8 are directed to a method of generating a pseudo-random number sequence, where a starting position of an array of digits is first selected, the digits are then regrouped with reference to the selected starting position so as to form a pseudo-random number.

2. Claims 9-48 are directed to an encoding / decoding method, where a key element is generated by performing an operation between each primary data element with a secondary data element.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/SG 99/00105

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member					
WO	96/42151	EP	872073	US	5613004	US	5687236
US	5276738	EP	614147				
EP	301383	JP	1036212	US	5901264	JP	1036213

END OF ANNEX